DOCUMENT RESUME

ED 233 077 TM 830 535

AUTHOR Lewis, Janice; Hoover, H. D.
TITLE Sex Differences on Standardized Academic Achievement Tests - A Longitudinal Study.
PUB DATE Apr 83
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; College Freshmen; High Schools; Language Tests; *Longitudinal Studies; Quantitative Tests; *Scores; *Sex Differences; Standardized Tests; Testing Programs
IDENTIFIERS *Iowa Testing Programs; Iowa Tests of Basic Skills; Iowa Tests of Educational Development

ABSTRACT

The records of graduates of Iowa high schools who entered the University of Iowa as freshmen in the fall of 1977 were secured in cooperation with the Iowa Testing Program. Test scores for grades 4, 6, and 8 on the Iowa Tests of Basic Skills and for grade 11 on the Iowa Tests of Educational Development were obtained for all students whose scores could be located in the records of the Iowa Testing Program. The data were grouped into five test areas: vocabulary, reading, language or expression, work-study or sources of information, and mathematics. The performance of males and females in these five areas was compared across grades. In this group of students who later entered college, girls consistently outscored boys in language or expression across the grades examined, but the boys' mean score was significantly higher than the girls' mean in mathematics at grade 11. The girls' mean score in mathematics was higher initially, but by sixth grade boys scored higher in mathematics on the average and continued to score higher than girls in grades 8 and 11. In the other three areas boys and girls scored similarly. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED233077

SEX DIFFERENCES ON STANDARDIZED ACADEMIC ACHIEVEMENT TESTS - A LONGITUDINAL STUDY

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

JANICE LEWIS and H. D. HOOVER

The University of Iowa

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. C. Lewis

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM830535

TM820 535

Sex Differences on Standardized Academic Achievement Tests: A Longitudinal Study

J. C. Lewis & H. D. Hoover, The University of Iowa

The records of graduates of Iowa high schools who entered the University of Iowa as freshmen in the fall of 1977 were secured in cooperation with the Iowa Testing Program. Test scores for grades 4, 6 and 8 on the Iowa Tests of Basic Skills and for grade 11 on the Iowa Tests of Educational Development were obtained for all students whose scores could be located in the records of the Iowa Testing Program. The data were grouped into five test areas: vocabulary, reading, language or expression, work-study or sources of information, and mathematics. The performance of males and females in these five areas was compared across grades. In this group of students who later entered college, girls consistently outscored boys in language or expression across the grades examined, but the boys' mean score was significantly higher than the girls' mean in mathematics at grade eleven. The girls' mean score in mathematics was higher initially, but by sixth grade boys scored higher in mathematics on the average and continued to score higher than girls in grades 8 and 11. In the other three areas boys and girls scored similarly.

Sex Differences on Standardized Academic Achievement Tests - A Longitudinal Study

JANICE LEWIS and H. D. HOOVER
The University of Iowa

Janice Lewis
332 Lindquist Center
The University of Iowa
Iowa City, IA 52242

Differences between males and females in academic achievement have been the focus of intensive discussion for many years. Lindquist and Stroud (1942) report sex differences in grades three through eight from the 1940 Iowa Every-Pupil Tests of Basic Skills. They find that girls are consistently superior in reading comprehension, vocabulary, work study skills, and language skills although the differences are small in reading comprehension after fifth grade. Small differences favoring boys are present for arithmetic in all grades, except in eighth grade where boys score significantly higher than girls. Maccoby and Jacklin (1974) discuss numerous research studies regarding sex differences in verbal and quantitative ability.

The debate continues in the 1980's, particularly in regard to sex differences in mathematics achievement. Benbow and Stanley (1980) report that in data collected during an eight-year period on gifted junior high school students, large sex differences are found in mathematical aptitude test scores favoring boys. They conclude that these differences probably do not result solely from socialization differences and may be due to superior male ability in mathematics. Based on data from the second mathematics assessment (1977-1978) by the National Assessment of Educational Progress and from the Women in Mathematics survey, Armstrong (1981) states that at age 13 females perform significantly better than males on computation but somewhat worse than males on problem-solving skills. By ages 17 and 18, males outscore females in all areas

TM 830 535

of mathematics assessment and are particularly favored on the applications or problem-solving subtest. Armstrong notes that these differences can not be explained entirely by differences in the number and level of mathematics courses taken. Fox (1981) also concludes that differences in exposure to mathematics instruction account for some but not all the discrepancy in achievement scores. Fox suggests that bias in achievement test items and differences in learning experiences outside school may be other explanations for sex differences on mathematics achievement tests. Meece, Parsons, Kaczala, Goff, and Futterman (1982) discuss possible sources of these sex differences. They conclude that sex differences on mathematics achievement tests do not consistently appear before tenth grade and that the differences are not large nor always present. However, they suggest that sex differences on mathematics achievement tests may appear earlier among gifted students than among students generally.

Sex differences in reading achievement have also received some scrutiny. Bank, Biddle, and Good (1980) note that numerous studies in the United States have shown that girls tend to outperform boys in reading and mathematics. They point out that girls read better than boys throughout the school years, but by adulthood no significant differences in reading ability are found between men and women. In mathematics the trend is just the opposite. No significant differences are found in early elementary school, but by adulthood there are large differences in mathematics achievement favoring men. Based upon cross-cultural studies, Sheridan and Fizdale believe that cultural

factors probably explain these differences, particularly in reading.

Some longitudinal studies of sex differences in achievement have been reported. Droege (1967) reports on a study by the United States Employment Service from 1958 to 1961 using the General Aptitude Test Battery (GATB) in which students were initially tested at ninth, tenth, or eleventh grade and retested at twelfth grade. Sex Differences in numerical and verbal aptitude are small. The mean score increases for verbal aptitude are not significantly different between males and females, but the mean score on the numerical aptitude subtest increases significantly more for males than females during the high school period. Hilton and Berglund (1974) report on a longitudinal study of mathematics achievement conducted by the Educational Testing Service during 1961 to 1967 in which students were tested in fifth, seventh, ninth, and eleventh grades using the math subtest of the Sequential Test of Educational Progress (STEP) and the quantitative subscale of the School and College Ability Test (SCAT). Hilton and Berglund find that males and females are approximately equal in mathematics achievement at the fifth grade level, but that males have successively higher mean scores at the higher levels. Hilton and Berglund compare students enrolled at the high school level in college preparatory programs to those in other programs such as vocational-technical, home economics, or business. On the STEP-Math, sex differences favoring males are significant at seventh, ninth, and eleventh grade for students in the college preparatory programs, while significant sex differences favoring males do not appear until eleventh grade in the other programs. When scores

from SCAT-Q are considered, sex differences favoring males are found at eleventh grade for all programs, but in the non-college group females outscore males at fifth grade. Hilton and Berglund also report on attitude data which was collected simultaneously. They find that at about ninth grade boys and girls begin to have increasingly different interests. Boys become more interested in science and mathematics relative to girls. As with achievement and ability, the sex differences in interests are larger in the college preparatory group than the non-college group. Brush (1980) also collected longitudinal data on attitudes regarding academic achievement both in mathematics and English from students in sixth through twelfth grade in 1976 to 1979. Brush reports that at sixth grade both boys and girls are generally neutral or slightly positive toward mathematics, but over time students of both sexes become less positive about mathematics. Boys consistently believe that math is easier than girls do and find math less anxiety-provoking.

The purpose of this study is to examine the differences in performance over time between boys and girls on standardized achievement tests. The achievement tests used in the study are the Iowa Tests of Educational Development (ITED) (Lindquist and Feldt, 1972) and the Iowa Tests of Basic Skills (ITBS) (Hieronymus and Lindquist, 1964, 1971). The ITED is a test battery constructed to measure general educational achievement in several content areas in grades nine through twelve. The ITBS is a multilevel battery which measures achievement in five content areas in grades three through nine.

METHOD

The records of 1,997 graduates of Iowa high schools who entered the University of Iowa as freshmen in the fall of 1977 were secured in cooperation with the Iowa Testing Programs. Test scores on the ITED and ITBS were obtained for grades 4, 6, 8, and 11 for all students whose scores could be located in the records of the Iowa Testing Programs. This data covered the years 1968-1975. The group included 959 males and 1038 females.

The variables used in this study were as follows: ITED achievement scores for grade 11 from the Vocabulary, Expression, Quantitative Thinking, and Sources of Information subtests as well as the Reading Total score and ITBS achievement scores for grades 4, 6, and 8 on the Vocabulary and Reading subtests and the Language Total, Work-Study Skills Total, and Mathematics Total scores. The Language Total score included subscores from Spelling, Punctuation, Capitalization, and Usage. The Mathematics Total score included subscores from Mathematics Concepts and Mathematics Problem-Solving. The Work-Study Total score included subscores from Map Reading, Reading Tables and Graphs, and Knowledge and Use of Reference Materials. The ITBS scores were grade equivalent scores, while the ITED scores were standard scores.

Summary statistics including means and standard deviations were computed using all available data on each variable for the entire group and for males and females separately. In order to follow exactly the same students throughout the time period, the group was reduced to include only students for whom there were complete records. This resulted in a group of 223 males and 234 females. Summary statistics

were computed for the reduced group. The data were grouped into five test areas: vocabulary, reading, language or expression, work-study or sources of information, and mathematics or quantitative thinking. An analysis of variance was performed on each of the 5 test areas using a split plot design (mixed design with one between and one within subject factor) in which subjects were crossed with respect to grade level and nested with respect to male-female category. Because of the different types of scores available for the ITBS and ITED, the scores were scaled within grade and test to a mean of 0.0 and a standard deviation of 1.0. Because of this transformation, any main effects due to grade were zero.

The ITBS scores for the entire group of 52,115 sixth grade students who took the ITBS during the 1970-1971 school year were summarized so that the average achievement level of the students in this study could be contrasted with an unselected group of sixth graders. The sixth grade was chosen for comparison because records categorized by sex were available for that year.

RESULTS AND DISCUSSION

The means, standard deviations, and n-counts for the entire group of incoming freshmen are shown in Table 1. In language skills the mean score for girls was higher than the mean for boys in all grades. In mathematics boys and girls were approximately equal at grade 4, but at all other grade levels the mean score for boys was larger than the mean for girls. Average performance on the vocabulary, reading, and work-study skills tests was similar for boys and girls in all grades.

The means and standard deviations for the group having complete

data (the reduced group) appear in Table 2. Both the original and the scaled means and standard deviations are shown. The means for the reduced group were somewhat higher than those for the entire incoming freshmen group for both boys and girls. The standard deviations for girls were smaller in the reduced data set than in the entire group, while some standard deviations for boys were larger in the reduced data set and others were larger in the data for the entire group. However, the patterns of male or female superiority were preserved in the reduced data set except for work-study skills. There the patterns of superiority were different, but the scores for boys and girls on work-study skills were similar at all grade levels in both groups.

The analysis of variance showed that neither the main effect for sex nor the sex by grade interaction was significant for vocabulary, reading, or work-study skills. For language, the sex main effect was highly significant ($F=47.29$, $p < .0001$), but the sex by grade interaction was not significant. However, for mathematics, only the sex by grade interaction was significant ($F=13.15$, $p < .0001$). Follow-up tests of this interaction showed that only at grade 11 was the difference between males' and females' mathematics mean score significant ($t=3.90$, $p < .001$). When boys and girls were considered separately, grade was found to be a significant factor in both cases ($F=7.37$, $p < .0001$ for girls; $F=5.93$, $p < .001$ for boys). The mathematics scores for girls could be divided into three groups which were significantly different from each other at $p < .05$; grade 4, grades 6 and 8, grade 11. For boys, the mathematics scores were divided into two groups: grade 4, and grades 6, 8, and 11. Furthermore, the scaled means in Table 2 showed that while the average mathematics score was higher for girls than boys at grade 4, the girls'

mean score declined at each grade level and the boys' mean increased at each grade.

When the mean male-female differences in mathematics were compared between grades, the comparisons between grade 4 and each of the other grades were significant at p .05 using a Scheffe standard of comparison. The comparison between grade 6 and grade 11 was also significant (Table 3). This suggests that female mathematics performance has already begun to fall behind male mathematics performance by grade 6.

The patterns present for mathematics and language skills, at least at grades 4, 6, and 8, are not likely to be caused by differential course taking. Most schools require courses in mathematics and language arts through grade 8. Very little choice of courses is available in most elementary schools, and even in grades 7 to 9, choices of mathematics and language arts courses are generally limited. At these grade levels, boys and girls should be receiving exposure to the same curricula in the mathematics and language areas. Despite the similarity of the curricula, girls clearly achieved higher scores in language arts than boys. In mathematics girls outscored boys at grade 4, but boys scored higher in the other three grades studied.

Compared to all students who took the ITBS in sixth grade during the 1970-1971 school year, the mean scores for both boys and girls in the reduced group ranked above the 75th percentile in every area except the boys' score in language which ranked at approximately the 72nd percentile. The high ranking of these students is an important factor in the sex differences that appeared. Table 4 shows the grade equivalent scores for boys and girls at given percentiles for Iowa sixth grade students in 1970. At the 10th, 25th, and 50th percentiles, girls

scored higher than boys in all areas. At the 75th percentile, girls scored slightly lower than boys in mathematics but had equal or higher scores in all other areas. However, at the 90th percentile, the boys' mean score was equal to or higher than the girls' mean in all areas except language skills. The ITBS Manual for School Administrators (Hieronymus, Lindquist, & Hoover, 1982) includes a table of sex differences which showed similar results. At the 10th, 25th, 50th, and 75th girls scored higher for most grade levels in all areas. At the 90th and 95th percentiles, boys achieved higher scores at most grade levels on the vocabulary test, the visual materials subtest of the work-study skills test, and the concepts and problem-solving subtests of the mathematics test. Girls continued to be superior in language skills, references, and computation even at the upper percentiles. Unpublished data (Forsyth, 1981) from the 1981 ITED scores of all Iowa high school juniors (Table 5) showed boys' scores to be higher than girls' scores in quantitative thinking at the 25th, 50th, and 75th percentiles, although the difference was largest at the 75th percentile. Girls earned higher scores in vocabulary, reading, sources of information, and expression. In summary, the ITBS and ITED data suggest that, had this been a lower achieving group, all sex differences except grade eleven mathematics probably would have favored girls.

Like Hilton and Berglund (1974), this study found that males outscored females in mathematics at eleventh grade in a group of students who later enter college. In this group, at grade 4 females were slightly superior to males in mathematics, but by grade 6 males were superior to females in mathematics and continued to be superior at the higher grade levels even though differential course-taking was not likely to be a

factor until grade 9 or 10. This change occurred at approximately the same age that Brush (1980) reported changes in girls' attitudes toward mathematics. This study did not support the assertion that girls outperform boys in verbal skills. Large differences were found favoring girls in language skills, but at least in this selected group, no significant differences were found between male and female mean scores for reading or vocabulary.

SUMMARY

This study compared the ITBS scores at fourth, sixth, and eighth grades and ITED scores at eleventh grade (1968-1975) of boys and girls who later enter college. These girls consistently outscored boys in language skills, but boys' mean scores were significantly higher than girls' scores in mathematics at grade eleven. In other areas boys and girls performed similarly. Girls were initially superior in mathematics, but by sixth grade boys began to outscore girls and continued to score higher than girls at the higher grade levels even though differential course-taking was not likely to be a factor at that grade level. Examination of data from ITBS and ITED suggested that the differences favoring boys in mathematics occurred mainly in the higher ability levels.

Table 1

Means and Standard Deviations by Sex
For the Incoming Freshmen at the University of Iowa (1977)

Achievement Area	Girls (n=1038)			Boys (n=959)		
	n	Mean	S.D.	n	Mean	S.D.
Vocabulary						
Grade 4	532	54.3	11.0	492	54.0	11.1
Grade 6	647	76.7	12.9	605	77.2	13.7
Grade 8	641	98.9	14.7	592	99.7	14.4
Grade 11	547	21.7	4.3	514	21.6	3.9
Reading						
Grade 4	531	55.9	11.5	492	55.1	12.4
Grade 6	647	76.9	12.1	605	77.0	13.1
Grade 8	640	100.3	13.8	592	99.5	15.1
Grade 11	543	22.5	4.5	512	22.4	4.8
Language						
Grade 4	531	57.4	9.8	493	52.5	11.1
Grade 6	647	81.1	12.4	606	74.5	14.1
Grade 8	641	101.4	15.4	591	94.2	16.6
Grade 11	548	21.4	3.7	514	19.7	4.0
Work-Study						
Grade 4	529	53.8	9.2	493	54.2	9.9
Grade 6	647	77.2	10.6	606	78.2	11.5
Grade 8	635	98.7	12.9	587	100.3	14.1
Grade 11	548	24.0	4.6	515	24.1	4.7
Mathematics						
Grade 4	527	51.4	8.0	491	51.3	9.0
Grade 6	647	73.6	10.8	605	75.5	11.4
Grade 8	635	96.5	14.2	587	99.4	14.5
Grade 11	548	21.2	5.7	515	23.4	5.6

Table 2

Means and Standard Deviations by Sex
For the Group with Complete Data

Achievement Area	Original				Scaled				Mean+ Difference Girls-Boys
	Girls (n=234)		Boys (n=223)		Girls (n=234)		Boys (n=223)		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
Vocabulary									
Grade 4	55.0	10.3	54.2	11.2	0.03	0.96	-0.04	1.04	0.07
Grade 6	77.1	12.0	77.7	13.9	-0.02	0.93	0.02	1.08	-0.05
Grade 8	99.9	13.7	100.3	13.5	-0.02	1.01	0.01	1.00	-0.03
Grade 11	21.7	4.2	21.6	3.9	0.02	1.03	-0.02	0.97	0.03
Reading									
Grade 4	56.9	10.6	55.6	12.5	0.05	0.92	-0.06	1.08	0.11
Grade 6	78.1	11.3	78.4	12.6	-0.01	0.95	0.01	1.06	-0.02
Grade 8	101.9	13.4	101.2	13.7	0.02	0.99	-0.03	1.01	0.05
Grade 11	22.6	4.2	22.6	4.6	0.00	0.95	0.00	1.05	0.00
Language									
Grade 4	59.1	9.3	53.2	11.1	0.26	0.88	-0.28	1.05	0.54**
Grade 6	83.2	11.0	75.5	13.4	0.29	0.86	-0.31	1.05	0.60**
Grade 8	103.6	13.1	95.3	15.5	0.27	0.88	-0.28	1.04	0.55**
Grade 11	21.5	3.5	19.8	3.9	0.22	0.92	-0.23	1.03	0.45**
Work-Study									
Grade 4	55.4	9.0	54.9	10.1	0.02	0.94	-0.03	1.06	0.05
Grade 6	79.1	9.8	79.2	11.4	-0.01	0.93	0.01	1.08	-0.01
Grade 8	101.0	11.1	100.9	14.2	0.00	0.87	0.00	1.12	0.00
Grade 11	24.4	4.1	24.5	4.6	-0.01	0.94	0.00	1.06	-0.02
Mathematics									
Grade 4	52.5	7.3	51.4	9.0	0.06	0.88	-0.07	1.10	0.13
Grade 6	74.9	10.3	76.3	11.4	-0.07	0.95	0.07	1.05	-0.14
Grade 8	98.1	12.8	100.0	14.3	-0.07	0.94	0.07	1.05	-0.14
Grade 11	21.6	5.4	23.5	5.4	-0.18	0.97	0.18	0.99	-0.36*

+positive differences favor girls
*significant $p < .05$
**significant $p < .001$

Table 3

t-tests Comparing Mean Male-Female Differences Between Grade

		Grade 4	Grade 6	Grade 8
Grade	6	6.26*	2.86*	2.74
	8	3.49*	0.09	
	11	3.40*		

* Significant at $p < .05$ using the Scheffe standard of comparison

Table 4

Grade Equivalent Scores at Given Percentiles

for Iowa Sixth Grade Students in 1970

	Percentiles									
	10th		25th		50th		75th		90th	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Vocabulary	37.6	39.5	48.7	50.5	61.2	61.9	74.5	74.5	85.1	84.8
Reading	41.8	45.6	50.4	53.7	61.7	64.3	74.3	75.4	84.2	84.5
Language Skills	37.4	44.6	45.5	54.7	57.7	68.4	71.6	81.7	83.6	90.5
Work Study Skills	42.5	45.6	51.1	54.7	63.5	65.9	75.4	75.9	84.3	83.5
Mathematics	43.2	44.3	51.3	53.2	62.8	63.7	73.3	73.0	82.4	81.0
Composite	42.6	46.0	50.5	54.3	61.5	64.9	72.8	75.1	81.7	82.9

Table 5

Standard Scores for Given Percentiles

for Iowa Eleventh Grade Students in 1981

	Percentiles					
	25th		50th		75th	
	Boys	Girls	Boys	Girls	Boys	Girls
Vocabulary	13.4	13.8	17.1	17.3	20.2	20.4
Reading Total	10.0	12.4	15.6	17.2	20.4	21.2
Expression	10.5	14.4	14.7	17.7	18.4	21.2
Sources of Information	11.0	13.2	17.3	19.1	23.0	23.7
Quantitative Thinking	10.8	10.4	16.0	15.0	21.3	19.5

REFERENCES

Armstrong, J. M. Achievement and participation of women in mathematics: Results of two national surveys. Journal for Research in Mathematics Education, 1981, 12, 356-372.

Bank, B. J., Biddle, B. J., & Good, T. L. Sex roles, classroom instruction, and reading achievement. Journal of Educational Psychology, 1980, 72, 119-132.

Benbow, C. P. & Stanley, J. C. Sex differences in mathematical ability: Fact or artifact? Science, 1980, 210, 1262-1264.

Brush, L. R. Encouraging Girls in Mathematics. Cambridge, Mass.: Abt Books, 1980.

Droege, R. C. Sex differences in aptitude maturation during high school. Journal of Counseling Psychology, 1967, 14, 407-411.

Forsyth, R. A. Unpublished data. Iowa City, Iowa: Iowa Testing Programs, 1981.

Fox, L. H. The Problem of Women and Mathematics. New York: Ford Foundation, 1981.

Hieronymus, A. N., Lindquist, E. F. Iowa Tests of Basic Skills, Forms 3 and 4. Boston: Houghton Mifflin, 1964.

Hieronymus, A. N., Lindquist, E. F. Iowa Tests of Basic Skills, Forms 5 and 6. Boston: Houghton Mifflin, 1971.

Hieronymus, A. N., Lindquist, E. F., & Hoover, H. D. Iowa Tests of Basic Skills Manual for School Administrators. Chicago: Riverside Publishing Co., 1982.

Hilton, T. L. & Berglund, G. W. Sex differences in mathematics achievement -- A longitudinal study. The Journal of Educational Research, 1974, 67, 231-237.

Lindquist, E. F. & Feldt, L. S. Iowa Tests of Educational Development, Grades 9-12, Forms X-6 and Y-6. Iowa City: The University of Iowa, 1972.

Maccoby, E. E. & Jacklin, C. N. Psychology of Sex Differences. Palo Alto, Calif.: Stanford University Press, 1974.

Meece, J. L., Parsons, J. E., Kaczala, C. M., Goff, S. B., & Futterman, R. Sex differences in math achievement: Toward a model of academic choice. Psychological Bulletin, 1982, 91, 324-348.

Sheridan, E. M. & Fizdale, B. Cultural factors and sex differences in reading and mathematics. ERIC, ED205936, 1981.

Stroud, J. B. & Lindquist, E. F. Sex differences in achievement in the elementary and secondary schools. Journal of Educational Psychology, 1942, 33, 657-667.